CLAIMS

What is claimed is:

1. A circuit arrangement for searching a parent code sequence for a target code sequence, comprising:

a shift register arrangement having a plurality of stages, wherein each stage stores a code of a subset of codes of the parent code sequence, and the shift register arrangement is adapted to periodically shift the subset of codes to form a new subset of codes with another code from the parent code sequence in a leading stage;

a matching circuit coupled to the shift register arrangement, the matching circuit adapted to ascertain code position matches between the subset of codes in the stages of the shift register arrangement and codes in corresponding code positions of the target code sequence, and provide a programmed binary value for each code position match; and

a pipelined adder arrangement coupled to the matching circuit, the adder arrangement adapted to sum the binary values for code position matches for each respective subset of codes.

2. The circuit arrangement of claim 1, wherein each stage of the shift register arrangement is adapted for storage of a code of character data.

3. The circuit arrangement of claim 1, wherein each stage of the shift register arrangement is adapted for storage of a code of a plurality of character data.

4. The circuit arrangement of claim 1, wherein the pipelined adder arrangement is a pipelined adder tree.

5. The circuit arrangement of claim 1, wherein the pipelined adder arrangement includes at least one stage of pipelined carry-save adders coupled to at least one stage of pipelined carry-propagate adders.

6. The circuit arrangement of claim 5, wherein the at least one stage of pipelined carry-save adders are adapted to provide a plurality of binary vectors responsive to the

quantity of code position matches, and the at least one stage of pipelined carry-propagate adders are adapted to add the plurality of binary vectors.

7. The circuit arrangement of claim 1, further comprising a pipelined summing circuit coupled to the pipelined adder arrangement and adapted to determine a moving sum of code position matches for a plurality of subsets of codes.

8. The circuit arrangement of claim 7, wherein the plurality of subsets of codes includes at least a most recent subset of codes and a next most recent subset of codes.

9. The circuit arrangement of claim 7, wherein the plurality of subsets of codes includes a first subset of codes and a prior subset of codes, wherein an intervening subset of codes is processed between the first subset of codes and the prior subset of codes.

10. The circuit arrangement of claim 1, wherein each subset of codes includes *n* contiguous codes from the parent code sequence.

11. The circuit arrangement of claim 1, wherein the matching circuit includes a plurality of programmable lookup tables, each lookup table having an input terminal coupled to an output terminal of a corresponding stage of the shift register arrangement and configured to provide a programmed value responsive to an input code value.

12. A method for searching a parent code sequence for a target code sequence, comprising:

shifting the parent code sequence through a shift register arrangement having a plurality of stages, wherein the shift register arrangement stores a subset of codes of the parent code sequence and each stage stores a code of the subset of codes, and each shift of the subset of codes forms a new subset of codes with another code from the parent code sequence in a leading stage;

determining in parallel whether the codes in the stages of the shift register arrangement are equal to codes of the target code sequence in corresponding code

positions, and generating in parallel signals of a programmed binary value for each equality of a subset code and a target code; and

summing the signals of the programmed binary value in a pipelined adder that generates a sum corresponding to each shift of the shift register arrangement.

13. The method of claim 12, further comprising:

determining, for each respective subset of codes, a probability of being the target code sequence as the sum of the binary values for code position matches for the respective subset of codes divided a total quantity of code positions in the target code sequence; and

associating the probability for each respective subset of codes with a unique identifier representative of a location within the parent code sequence at which the respective subset of codes exists.

14. The method of claim 12, wherein the parent code sequence represents a genome.

15. The method of claim 14, wherein each code of the parent code sequence is representative of a nucleotide type.

16. The method of claim 15, wherein the nucleotide type is selected from the group consisting of: adenine, thymine, guanine, and cytosine.

17. The method of claim 14, wherein the genome is a human genome.

18. The method of claim 12, further comprising:

configuring a plurality of lookup tables to generate respective signals of the programmed binary value when addressed by codes equal to codes of the target code sequence; and

addressing the lookup tables with the codes of the subset of codes.

19. The method of claim 12, further comprising generating a moving sum, for n subsets of codes, of sums of the signals of the selected binary value.

20. The method of claim 19, wherein the *n* subsets of codes includes at least a most recent subset of codes and a next most recent subset of codes.

21. The method of claim 19, wherein the *n* subsets of codes includes a first subset of codes and a prior subset of codes, wherein an intervening subset of codes is processed between the first subset of codes and the prior subset of codes.

22. The method of claim 12, wherein each subset of codes includes *m* contiguous codes from the parent code sequence.

23. An apparatus for searching a parent code sequence for a target code sequence, each code in the parent code sequence having a parent-relative position, comprising:

means for periodically selecting subsets of codes of the parent code sequence, each code in the subset having a relative subset-code position defined by the parent-relative position, and each subset of codes differing from other subsets by parent-relative positions of the codes in the subset;

means for determining in parallel whether each code at a subset-code position in a subset of codes is equal to a code of the target code sequence in a corresponding target-code position, and generating in parallel signals of a selected binary value for each equality of a subset code and the target code; and

means for summing the signals of the selected binary value.